

generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequence is typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) and a hybridization probe were synthesized:

forward PCR primer 1 CAACGTGATTTCAAAGCTGGGCTC (SEQ ID NO:517)

forward PCR primer 2 GCCTCGTATCAAGAATTTCC (SEQ ID NO:518)

forward PCR primer 3 AGTGGAAGTCGACCTCCC (SEQ ID NO:519)

reverse PCR primer 1 CTCACCTGAAATCTCTCATAGCCC (SEQ ID NO:520)

hybridization probe 1 CGCAAAACCCATTTTGGGAGCAGGAATCCAATCATGTCTGTGATGGTGG (SEQ ID NO:521)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO298 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25). The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO298 (herein designated UNQ261 [DNA39975-1210]) (SEQ ID NO:514), and the derived protein sequence for PRO298 (SEQ ID NO:515).

The entire nucleotide sequence of UNQ261 (DNA39975-1210) is shown in Figure 218 (SEQ ID NO:514). Clone DNA39975-1210 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 375-377. The predicted polypeptide precursor is 364 amino acids long. The protein contains four putative transmembrane domains between amino acid positions 36-55 (type II TM), 65-84, 188-208, and 229-245, respectively. A putative N-linked glycosylation site starts at amino acid position 253. In addition, the following features have been identified in the protein sequence: cAMP- and cGMP-dependent protein kinase phosphorylation site, starting at position 8; N-myristoylation sites starting a position 173 and 262, respectively; and a ZP domain between amino acid positions 45-60. Clone DNA39975-1210 has been deposited with ATCC (April 21, 1998) and is assigned ATCC deposit no.209783.

#### EXAMPLE 96: Isolation of cDNA Clones Encoding Human PRO337

A cDNA sequence identified in the amylase screen described in Example 2 above is herein designated DNA42301 (Figure 223, SEQ ID NO:524). The DNA42301 sequence was then compared to other EST

sequences using phrap as described in Example 1 above and a consensus sequence designated herein as DNA28761 was identified. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO337 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of DNA43316-1237 is shown in Figure 221 (SEQ ID NO:522). Clone DNA43316-1237 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 (Figure 221; SEQ ID NO:522). The predicted polypeptide precursor is 344 amino acids long. Clone DNA43316-1237 has been deposited with ATCC and is assigned ATCC deposit no. 209487

Based on a BLAST-2 and FastA sequence alignment analysis of the full-length sequence, PRO337 shows amino acid sequence identity to rat neurotrimin (97%).

#### EXAMPLE 97: Isolation of cDNA Clones Encoding Human PRO403

##### Introduction:

Human thrombopoietin (THPO) is a glycosylated hormone of 352 amino acids consisting of two domains. The N-terminal domain, sharing 50% similarity to erythropoietin, is responsible for the biological activity. The C-terminal region is required for secretion. The gene for thrombopoietin (THPO) maps to human chromosome 3q27-q28 where the six exons of this gene span 7 kilobase base pairs of genomic DNA (Chang et al., Genomics 26: 636-7 (1995); Foster et al., Proc. Natl. Acad. Sci. USA 91: 13023-7 (1994); Gurney et al., Blood 85: 981-988 (1995). In order to determine whether there were any genes encoding THPO homologues located in close proximity to THPO, genomic DNA fragments from this region were identified and sequenced. Three P1 clones and one PAC clones (Genome Systems Inc., St. Louis, MO; cat. Nos. P1-2535 and PAC-6539) encompassing the THPO locus were isolated and a 140 kb region was sequenced using the ordered shotgun strategy (Chen et al., Genomics 17: 651-656 (1993)), coupled with a PCR-based gap filling approach. Analysis reveals that the region is gene-rich with four additional genes located very close to THPO: tumor necrosis factor-receptor type 1 associated protein 2 (TRAP2) and elongation initiation factor gamma (eIF4), chloride channel 2 (CLCN2) and RNA polymerase II subunit hRPB17. While no THPO homolog was found in the region, four novel genes have been predicted by computer-assisted gene detection (GRAIL)(Xu et al., Gen. Engin. 16: 241-253 (1994), the presence of CpG islands (Cross, S. and Bird, A., Curr. Opin. Genet. & Devel. 5: 109-314 (1995), and homology to known genes (as detected by WU-BLAST2.0)(Altschul and Gish, Methods Enzymol. 266: 460-480 (1996) (<http://blast.wustl.edu/blast/README.html>).

##### Procedures:

##### P1 and PAC clones:

The initial human P1 clone was isolated from a genomic P1 library (Genome Systems Inc., St. Louis, MO; cat. no.: P1-2535) screened with PCR primers designed from the THPO genomic sequence (A.L. Gurney, et al., Blood 85: 981-88 (1995). PCR primers were designed from the end sequences derived from this P1 clone

were then used to screen P1 and PAC libraries (Genome Systems, Cat. Nos.: P1-2535 & PAC-6539) to identify overlapping clones (PAC1, p1.t, and P1.u). The 3'-end sequence from PAC.z was used to define the primers used for the screening of a human BAC library (Genome Systems Inc., St. Louis, MO; Cat. No.: BDTW-4533A).

#### Ordered Shotgun Strategy:

The Ordered Shotgun Strategy (OSS) (Chen et al., Genomics 17: 651-656 (1993)) involves the mapping and sequencing of large genomic DNA clones with a hierarchical approach. The P1 or PAC clone was sonicated and the fragments subcloned into lambda vector ( $\lambda$ Bluestar) (Novagen, Inc., Madison, WI; cat. no. 69242-3). The lambda subclone inserts were isolated by long-range PCR (Barnes, W. Proc. Natl. Acad. Sci. USA 91: 2216-2220 (1994) and the ends sequenced. The lambda-end sequences were overlapped to create a partial map of the original clone. Those lambda clones with overlapping end-sequences were identified, the insets subcloned into a plasmid vector (pUC18 or pUC19, Hoefer Pharmacia Biotech, Inc., San Francisco, CA, Cat. Nos. 27-4949-01 and 27-4951-01) and the ends of the plasmid subclones were sequenced and assembled to generate a contiguous sequence. This directed sequencing strategy minimizes the redundancy required while allowing one to scan for and concentrate on interesting regions.

In order to define better the THPO locus and to search for other genes related to the hematopoietin family, five genomic clones were isolated from this region by PCR screening of human P1 and PAC libraries (Genome System, Inc., Cat. Nos.: P1-2535 and PAC-6539).

The sizes of the genomic fragments are as follows: P1.t is 40 kb; P1.g is 70 kb; P1.u is 70 kb; PAC.z is 200 kb; and BAC.1 is 80 kb. Approximately 75% (140 kb) of the 190 kb genomic DNA region was sequenced by the Ordered Shotgun Strategy (OSS) (Chen et al., Genomics 17: 651-56 (1993), and assembled into contigs using AutoAssemblerTM (Applied Biosystems, Perkin Elmer, Foster City, CA, cat. no. 903227). The preliminary order of these contigs was determined by manual analysis. There were 47 contigs the 140 kb region. A PCR-based approach to ordering the contigs and filling in the gaps was employed. The following summarizes the number and sizes of the gaps. The 50 kb of sequence unique to BAC.1 was sequenced by a total shotgun approach with a ten-fold redundancy.

<u>Size of gap</u>	<u>number</u>
< 50 bp	13
50-150 bp	7
150-300 bp	7
300-1000 bp	10
1000-5000 bp	7
> 5000 bp	2 ((15,000 bp)

#### DNA sequencing:

ABI DYE-primerTM chemistry (PE Applied Biosystems, Foster City, CA; Cat. No.: 402112) was used to end-sequence the lambda and plasmid subclones. ABI DYE-terminaterTM chemistry (PE Applied Biosystems, Foster City, CA, Cat. No: 403044) was used to sequence the PCR products with their respective PCR primers. The sequences were collected with an ABI377 instrument. For PCR products larger than 1kb, walking primers